ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/380861219

Outsourcing eureka moments to artificial intelligence

Article *in* Nature Computational Science · May 2024

DOI: 10.1038/s43588-024-00633-4

CITATIONS
0

READS
48

1 author:

Martijn Meeter
Vrije Universiteit Amsterdam
239 PUBLICATIONS **6,822** CITATIONS

SEE PROFILE

All content following this page was uploaded by Martijn Meeter on 25 October 2024.

The user has requested enhancement of the downloaded file.

Outsourcing eureka moments to artificial intelligence

Martijn Meeter
LEARN! Research Institute
Vrije Universiteit Amsterdam
m.meeter@vu.nl

Last draft before submission to Nat Comput Sci

Can one discover the laws of human behavior using Artificial Intelligence? A new study shows that this is a may be possible. Recently, The Economist dove into the growing role of artificial intelligence (AI) in science. AI has been increasingly used by scientists not just for number crunching but also for the essence of research: scientific discovery [1]. The new study in *Nature Computational Science* by Sannyuya Liu and colleagues [2] provides perhaps the first example of scientific discovery by AI in the field of cognitive psychology.

Liu and colleagues l aimed to discover new laws governing learning from the bottom up, by training machine learning algorithms on a large data set. In particular, they analyzed the data of 6000 individuals[3] who tried to master a range of cognitive skills (e.g., math, but also a better memory) by doing exercises on an online learning platform, Luminosity®.

In the first step, the authors predicted whether an individual would correctly perform the next exercises using a deep learning network. Deep learning networks are versatile pattern detectors that are also at the core of large language models. One downside of these models is that they function as a black box: once they reach their ability to predict the next sample in the data, it is very hard to discern on the basis of what they make these predictions. To counter this, the authors used a second algorithm, symbolic regression, to translate the deep learning network's trained weights into symbolic laws that can be interpreted by humans. Figure 1 gives an example of how the two algorithms interact: at the lowest rung is the data. A deep neural network was used to predict, based on previous sessions, the outcome of the next session. After these predictions were optimized, the symbolic regression algorithm extracted mathematical formulas that capture the learning performed by the deep neural network. In Figure 1, for example, a learning law where mastery is a function of the logarithm of the number of sessions is depicted. For the six skills they tested, their method was able to extract learning laws that equaled or outperformed the classic ones proposed in the literature [4]. Some of the laws had a completely new mathematical form (in particular the logarithmic one), others introduced new interactions between the different skills during learning.

While it is clear that the algorithms can offer formulas that fit the data well, it is still unclear whether they truly capture human learning, mainly for two reasons. First, the study does not convincingly show that the extracted laws are the ones that generated the data. While the authors tested their algorithm on simulated data, they did not simulate true raw data, but used a smooth transformation of the data as input to their model (concretely, they used mastery estimates that result from applying item-response theory, a well-established psychometric framework that was also used in their data fits). In this way, they abstracted away from the full noisiness of human data, and thus makes the job of the algorithms artificially easy.

The second reason is that even if the learning laws found by the algorithms would be the ones underlying Luminosity® data, it is unclear whether they generalize to other tasks. As an example, it is not clear, for students learning topography in a classroom, whether or not their progress would be described by any of the learning laws proposed by the authors. This is unknown, and in fact it is rather unlikely, given that the different skills investigated by Liu and colleagues yielded different learning laws. This implies that none of these laws generalized to any of the other skills. However, that is not a problem of the AI algorithms, but rather, a weakness of the underlying paradigm of trying to fit formulas to data [5] – also known as curve fitting.

While one can quibble about the particular application presented here, their work is still very relevant in how interpretable results are extracted from a data set using AI algorithms. If there are

mathematical laws to be found in the chaos of human behavior, this study gives some confidence that we can outsource their discovery to AI techniques. Whether there any such laws to be discovered is, alas, an open question.

1 The Economist. AI Science. Issue of 16 September 2023 (London, 2023).

2 Liu, S., Li, Q., Shen, X., Sun, J. & Yang, Z. Automated discovery of symbolic laws governing skill acquisition from naturally occurring data. *Nature Computational Science* **in press** (2024).

3 Steyvers, M. & Schafer, R. J. Inferring latent learning factors in large-scale cognitive training data. *Nature Human Behaviour* **4**, 1145-1155 (2020). https://doi.org:10.1038/s41562-020-00935-3

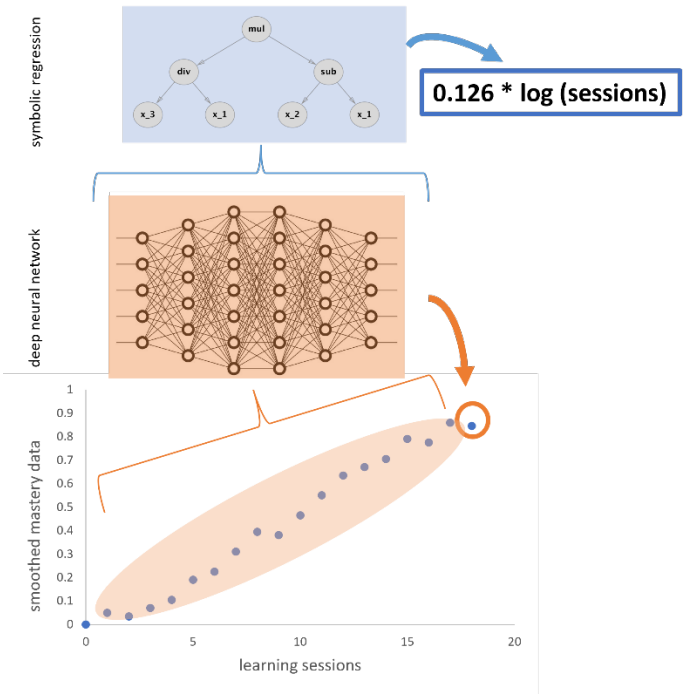
4 Murre, J. M. J. How averaging individual curves transforms their shape: Mathematical analyses with application to learning and forgetting curves. *Journal of Mathematical Psychology* **117**, 102816 (2023). https://doi.org:https://doi.org/10.1016/j.jmp.2023.102816

5 Meeter, M., Jehee, J. & Murre, J. Neural models that convince: Model hierarchies and other strategies to bridge the gap between behavior and the brain. *Phil Psychol* **20**, 749-772 (2007). https://doi.org:10.1080/09515080701694128

Competing interests: None.

Figure 1

AI techniques used by Liu and colleagues to extract learning laws from Luminosity® data. Bottom level: the data was aggregated at session level, and analyzed with item-response theory to yield a continuous mastery estimate. Middle level: a deep neural network tried to predict the data from the next session from data of the previous sessions, both from the skill that was trained in the predicted session as well as sessions from other skills. Top level: a symbolic regression algorithm then extracted a mathematical approximation of the weights learned by the deep neural network, thus capturing how the network made its predictions. In this case, this results in a logarithmic “learning law”.



View publication stats